

## SEQUENCE LISTING

&lt;110&gt; CNRS

<120> STARCH GRANULES CONTAINING A RECOMBINANT POLYPEPTIDE  
OF INTEREST, A METHOD OF OBTAINING THEM AND THEIR  
USES

&lt;130&gt; WOB 99AB CNR AMYL

&lt;140&gt;

&lt;141&gt;

&lt;150&gt; FR9906494

&lt;151&gt; 1999-05-21:

&lt;160&gt; 9

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 3117

&lt;212&gt; DNA

&lt;213&gt; Chlamydomonas reinhardtii

&lt;400&gt; 1

```

gccaggcgga gcgtatggct gttgcctcta ccagccgccc cagcagcgcg cgtcctatcg 60
tcatcaatgc cgcgtcggtc ggtgtcaaga agaccgcgaa ccagctgctg cgtgagcttg 120
ctcgtggctc cgcacgcaag tccacctcgc gctcggctgt tactggtgcc actggtgcc 180
cttgcgcgct ggacatcgct atggttgctg ctgaggtcgc cccttgggtcc aagacgggcg 240
gcctgggcga tgtgactggg ggcctgccta ttgagctggg caagcgcggc caccgcgtca 300
tgaccattgc ccctcgctac gaccagtaag ctgacgcctg ggacacctcg gtggtcgtgg 360
acatcatggg cgagaaggtc cgctaacttc actccatcaa gaagggcgtg caccgcgtgt 420
ggattgacca cccctgggtc ctggccaagg tctggggcaa gaccggctcc aagctgtacg 480
gcccccgctc cggcgctgac tacctggaca accacaagcg cttcgccctg ttctgcaagg 540
ccgctattga ggctgcccgc gtgctgcctc tcggccccgg cgaggactgc gtcttcgtgg 600
ccaacgactg gcaactccgc ctggtgcccg tcctgctgaa ggacgagtac cagcccaagg 660
gccagttcac caaggccaag tcggtgctgg ctatccacaa catcgcttc cagggccgca 720
tgtgggagga ggctttcaag gacacgaagc tgcccaggc cgcctttgac aagctggcct 780
tctcggacgg ctatgccaa gtttacactg aggccacccc catggaggag gacgagaagc 840
ccccgctgac gggaaagacc tacaagaaga tcaactggct gaagggtggc attatcgccg 900
ccgacaagct ggtgactgtg tcgcccact acgcgaccga gatcgctgcc gatgccgccg 960
gcggtgtgga gctggacacc gtcacccgcg ccaagggcatt tgagggcatt gtgaacggca 1020
tggacattga ggagtgaac cccaagaccg acaagttcct gtctgcgccg tacgaccaga 1080
acagcgtcta cgcggcaag gccgcggcca aggaggccct gcaggccgag ctgggcctgc 1140
ctgtggaccc caccgcccc ctgttcgcct tcatcgcccg cctggaggag cagaagggtg 1200
tggacatcat cctggccgcc ctgcccaga tcctggccac cccaagggtg cagatcgcca 1260
tcctgggtac cggcaaggcc gcctacgaga agctggtgaa cgccatcggc accaagtaca 1320
agggccgcgc caaggcgctg gtcaagttct cggcgccctt ggcgcacatg ctaccgccg 1380
gcgcccactt catgctggtg ccctcgcgct tcgagccctg cggcctgatc cagctgcacg 1440
ccatgacta cggtaccgtg ccctggttag cctccaccgg cggcctggtc gacaccgtca 1500
aggaggcgt caccggcttc cacatgggag ccctgaaccc cgacaagctg gacgaggctg 1560
acgccgagcg cctggccgcc accgtgcgac gtgccagcga ggtgtttgag ggcggccgct 1620
accccgagat ggtggccaac tgcatacgcc aggcctgtc ctggtccaag cccgccaga 1680
agtgggaggg cctgctggag gaggtggtgt accgcaaggg cggcgtggcc accgccaaga 1740
aggaggagat caaggtgccc gttgccgaga agatccccgg cgacctgccc gccgtgtcct 1800
acgcccccaa caccctgaag ccctgtgctg cctccgtgga gggcaacggc gccgccgcgc 1860
ccaaggtcgg caccaccgcc ccgcctatgg gcgctggcg cgcgaccacc ccctcgggcc 1920
cctcggccgc cgccgccacc cccaaggtga ccacctaaa gcccgccctg cccgccaccg 1980
ccaagcccaa gaccgctggc ctcaagctgg ccggtgaggg ctccaccacc tcgacctcgg 2040
agaacggcgc tgcctccaac ggcaacggca acggtgcctc ggcctccaag acctcggctg 2100

```

09/980771  
 JC19  
 d PCT/PTO 15 NOV 2001

```
<210> 2
<211> 2124
<212> DNA
<213> Artificial sequence
```

```
<220>
<221> CDS
<222> (1)..(2124)
```

<400>	2																
atg gct gtt gcc tct acc agc cgc ccc agc agc gcg cgt cct atc gtc																	48
Met Ala Val Ala Ser Thr Ser Arg Pro Ser Ser Ala Arg Pro Ile Val																	
1				5					10					15			
atc aat gcc gcg tcg ttc ggt gtc aag aag acc gcg aac cag ctg ctg																	96
Ile Asn Ala Ala Ser Phe Gly Val Lys Lys Thr Ala Asn Gln Leu Leu																	
		20					25						30				
cgt gag ctt gct cgt ggc tcc gca cgc aag tcc acc tcg cgc tcg gct																	144
Arg Glu Leu Ala Arg Gly Ser Ala Arg Lys Ser Thr Ser Arg Ser Ala																	
		35					40					45					
gtt act ggt gcc act ggt gcc act tgc gcg ctg gac atc gtg atg gtt																	192
Val Thr Gly Ala Thr Gly Ala Thr Cys Ala Leu Asp Ile Val Met Val																	
	50					55					60						
gct gct gag gtc gcc cct tgg tcc aag acg ggc ggc ctg ggc gat gtg																	240
Ala Ala Glu Val Ala Pro Trp Ser Lys Thr Gly Gly Leu Gly Asp Val																	
65					70				75						80		
act ggt ggc ctg cct att gag ctg gtc aag cgc ggc cac cgc gtc atg																	288
Thr Gly Gly Leu Pro Ile Glu Leu Val Lys Arg Gly His Arg Val Met																	
				85				90						95			
acc att gcc cct cgc tac gac cag tac gct gac gcc tgg gac acc tcg																	336
Thr Ile Ala Pro Arg Tyr Asp Gln Tyr Ala Asp Ala Trp Asp Thr Ser																	
		100					105						110				

gtg Val	gtc Val	gtg Val	gac Asp	atc Ile	atg Met	ggc Gly	gag Glu	aag Lys	gtc Val	cgc Arg	tac Tyr	ttc Phe	cac His	tcc Ser	atc Ile	384
115																
aag Lys	aag Lys	ggc Gly	gtg Val	cac His	cgc Arg	gtg Val	tgg Trp	att Ile	gac Asp	cac His	ccc Pro	tgg Trp	ttc Phe	ctg Leu	gcc Ala	432
130																
aag Lys	gtc Val	tgg Trp	ggc Gly	aag Lys	acc Thr	ggc Gly	tcc Ser	aag Lys	ctg Leu	tac Tyr	ggc Gly	ccc Pro	cgc Arg	tcc Ser	ggc Gly	480
145																
gct Ala	gac Asp	tac Tyr	ctg Leu	gac Asp	aac Asn	cac His	aag Lys	cgc Arg	ttc Phe	gcc Ala	ctg Leu	ttc Phe	tgc Cys	aag Lys	gcc Ala	528
165																
gct Ala	att Ile	gag Glu	gct Ala	gcc Ala	cgc Arg	gtg Val	ctg Leu	ccc Pro	ttc Phe	ggc Gly	ccc Pro	ggc Gly	gag Glu	gac Asp	tgc Cys	576
180																
gtc Val	ttc Phe	gtg Val	gcc Ala	aac Asn	gac Asp	tgg Trp	cac His	tcc Ser	gcc Ala	ctg Leu	gtg Val	ccc Pro	gtc Val	ctg Leu	ctg Leu	624
195																
aag Lys	gac Asp	gag Glu	tac Tyr	cag Gln	ccc Pro	aag Lys	ggc Gly	cag Gln	ttc Phe	acc Thr	aag Lys	gcc Ala	aag Lys	tcg Ser	gtg Val	672
210																
ctg Leu	gct Ala	atc Ile	cac His	aac Asn	atc Ile	gcc Ala	ttc Phe	cag Gln	ggc Gly	cgc Arg	atg Met	tgg Trp	gag Glu	gag Glu	gct Ala	720
225																
ttc Phe	aag Lys	gac Asp	acg Thr	aag Lys	ctg Leu	ccc Pro	cag Gln	gcc Ala	gcc Ala	ttt Phe	gac Asp	aag Lys	ctg Leu	gcc Ala	ttc Phe	768
245																
tcg Ser	gac Asp	ggc Gly	tat Tyr	gcc Ala	aag Lys	gtt Val	tac Tyr	act Thr	gag Glu	gcc Ala	acc Thr	ccc Pro	atg Met	gag Glu	gag Glu	816
260																
gac Asp	gag Glu	aag Lys	ccc Pro	ccg Pro	ctg Leu	acg Thr	gga Gly	aag Lys	acc Thr	tac Tyr	aag Lys	aag Lys	atc Ile	aac Asn	tgg Trp	864
275																
ctg Leu	aag Lys	ggc Gly	ggc Gly	att Ile	atc Ile	gcc Ala	gcc Ala	gac Asp	aag Lys	ctg Leu	gtg Val	act Thr	gtg Val	tcg Ser	ccc Pro	912
290																
aac Asn	tac Tyr	gcg Ala	acc Thr	gag Glu	atc Ile	gct Ala	gcc Ala	gat Asp	gcc Ala	gcc Ala	ggc Gly	ggt Gly	gtg Val	gag Glu	ctg Leu	960
305																
gac Asp	acc Thr	gtc Val	atc Ile	cgc Arg	gcc Ala	aag Lys	ggc Gly	att Ile	gag Glu	ggc Gly	att Ile	gtg Val	aac Asn	ggc Gly	atg Met	1008
325																
gac Asp	att Ile	gag Glu	gag Glu	tgg Trp	aac Asn	ccc Pro	aag Lys	acc Thr	gac Asp	aag Lys	ttc Phe	ctg Leu	tct Ser	gcg Ala	ccc Pro	1056
340																

tac gac cag aac agc gtc tac gcc ggc aag gcc gcc gcc aag gag gcc	1104
Tyr Asp Gln Asn Ser Val Tyr Ala Gly Lys Ala Ala Ala Lys Glu Ala	
355 360 365	
ctg cag gcc gag ctg ggc ctg cct gtg gac ccc acc gcc ccc ctg ttc	1152
Leu Gln Ala Glu Leu Gly Leu Pro Val Asp Pro Thr Ala Pro Leu Phe	
370 375 380	
gcc ttc atc ggc cgc ctg gag gag cag aag ggt gtg gac atc atc ctg	1200
Ala Phe Ile Gly Arg Leu Glu Glu Gln Lys Gly Val Asp Ile Ile Leu	
385 390 395 400	
gcc gcc ctg ccc aag atc ctg gcc acc ccc aag gtg cag atc gcc atc	1248
Ala Ala Leu Pro Lys Ile Leu Ala Thr Pro Lys Val Gln Ile Ala Ile	
405 410 415	
ctg ggt acc ggc aag gcc gcc tac gag aag ctg gtg aac gcc atc ggc	1296
Leu Gly Thr Gly Lys Ala Ala Tyr Glu Lys Leu Val Asn Ala Ile Gly	
420 425 430	
acc aag tac aag ggc cgc gcc aag ggc gtg gtc aag ttc tcg gcg ccc	1344
Thr Lys Tyr Lys Gly Arg Ala Lys Gly Val Val Lys Phe Ser Ala Pro	
435 440 445	
ctg gcg cac atg ctc acc gcc ggc gcc gac ttc atg ctg gtg ccc tcg	1392
Leu Ala His Met Leu Thr Ala Gly Ala Asp Phe Met Leu Val Pro Ser	
450 455 460	
cgc ttc gag ccc tgc ggc ctg atc cag ctg cac gcc atg cac tac ggt	1440
Arg Phe Glu Pro Cys Gly Leu Ile Gln Leu His Ala Met His Tyr Gly	
465 470 475 480	
acc gtg ccc gtg gta gcc tcc acc ggc ggc ctg gtc gac acc gtc aag	1488
Thr Val Pro Val Val Ala Ser Thr Gly Gly Leu Val Asp Thr Val Lys	
485 490 495	
gag ggc gtc acc ggc ttc cac atg ggc gcc ctg aac ccc gac aag ctg	1536
Glu Gly Val Thr Gly Phe His Met Gly Ala Leu Asn Pro Asp Lys Leu	
500 505 510	
gac gag gct gac gcc gac gcc ctg gcc gcc acc gtg cgc cgt gcc agc	1584
Asp Glu Ala Asp Ala Asp Ala Leu Ala Ala Thr Val Arg Arg Ala Ser	
515 520 525	
gag gtg ttt gcg ggc ggc cgc tac ccc gag atg gtg gcc aac tgc atc	1632
Glu Val Phe Ala Gly Gly Arg Tyr Pro Glu Met Val Ala Asn Cys Ile	
530 535 540	
agc cag gac ctg tcc tgg tcc aag ccc gcc cag aag tgg gag ggc ctg	1680
Ser Gln Asp Leu Ser Trp Ser Lys Pro Ala Gln Lys Trp Glu Gly Leu	
545 550 555 560	
ctg gag gag gtg gtg tac ggc aag ggc ggc gtg gcc acc gcc aag aag	1728
Leu Glu Glu Val Val Tyr Gly Lys Gly Val Ala Thr Ala Lys Lys	
565 570 575	
gag gag atc aag gtg ccc gtt gcc gag aag atc ccc ggc gac ctg ccc	1776
Glu Glu Ile Lys Val Pro Val Ala Glu Lys Ile Pro Gly Asp Leu Pro	
580 585 590	

gcc	gtg	tcc	tac	gcc	ccc	aac	acc	ctg	aag	ccc	gtg	tcc	gcc	tcc	gtg	1824
Ala	Val	Ser	Tyr	Ala	Pro	Asn	Thr	Leu	Lys	Pro	Val	Ser	Ala	Ser	Val	
		595					600					605				
gag	ggc	aac	ggc	gcc	gcc	gcg	ccc	aag	gtc	ggc	acc	acc	gcc	ccc	gcc	1872
Glu	Gly	Asn	Gly	Ala	Ala	Ala	Pro	Lys	Val	Gly	Thr	Thr	Ala	Pro	Ala	
	610					615					620					
atg	ggc	gcg	tgg	cgc	gcg	acc	acc	ccc	tcg	ggc	ccc	tcg	ccc	gcc	gcc	1920
Met	Gly	Ala	Trp	Arg	Ala	Thr	Thr	Pro	Ser	Gly	Pro	Ser	Pro	Ala	Ala	
625					630					635					640	
gcc	acc	ccc	aag	gtg	acc	acc	tac	aag	ccc	gcc	ctg	ccc	gcc	acc	gcc	1968
Ala	Thr	Pro	Lys	Val	Thr	Thr	Tyr	Lys	Pro	Ala	Leu	Pro	Ala	Thr	Ala	
				645					650					655		
aag	ccc	aag	acc	gct	ggc	ctc	aag	ctg	gcc	ggt	gag	gcc	tcc	acc	acc	2016
Lys	Pro	Lys	Thr	Ala	Gly	Leu	Lys	Leu	Ala	Gly	Glu	Ala	Ser	Thr	Thr	
			660					665					670			
tcg	acc	tcg	gag	aac	ggc	gct	gcc	tcc	aac	ggc	aac	ggc	aac	ggt	gcc	2064
Ser	Thr	Ser	Glu	Asn	Gly	Ala	Ala	Ser	Asn	Gly	Asn	Gly	Asn	Gly	Ala	
			675				680					685				
tcg	gcc	tcc	aag	acc	tcg	gct	gcc	aag	ccc	ctg	gtc	tcc	gcc	gcc	acc	2112
Ser	Ala	Ser	Lys	Thr	Ser	Ala	Ala	Lys	Pro	Leu	Val	Ser	Ala	Ala	Thr	
	690					695					700					
cgc	aag	tcc	gcc													2124
Arg	Lys	Ser	Ala													
705																
<210> 3																
<211> 708																
<212> PRT																
<213> Artificial sequence																
<223> Description of the artificial sequence: fragment																
of the complete sequence of cDNA coding for the GBSSI																
of Chlamydomonas reinhardtii																
<400> 3																
Met	Ala	Val	Ala	Ser	Thr	Ser	Arg	Pro	Ser	Ser	Ala	Arg	Pro	Ile	Val	
1				5					10					15		
Ile	Asn	Ala	Ala	Ser	Phe	Gly	Val	Lys	Lys	Thr	Ala	Asn	Gln	Leu	Leu	

Thr Ile Ala Pro Arg Tyr Asp Gln Tyr Ala Asp Ala Trp Asp Thr Ser  
 100 105 110  
 Val Val Val Asp Ile Met Gly Glu Lys Val Arg Tyr Phe His Ser Ile  
 115 120 125  
 Lys Lys Gly Val His Arg Val Trp Ile Asp His Pro Trp Phe Leu Ala  
 130 135 140  
 Lys Val Trp Gly Lys Thr Gly Ser Lys Leu Tyr Gly Pro Arg Ser Gly  
 145 150 155 160  
 Ala Asp Tyr Leu Asp Asn His Lys Arg Phe Ala Leu Phe Cys Lys Ala  
 165 170 175  
 Ala Ile Glu Ala Ala Arg Val Leu Pro Phe Gly Pro Gly Glu Asp Cys  
 180 185 190  
 Val Phe Val Ala Asn Asp Trp His Ser Ala Leu Val Pro Val Leu Leu  
 195 200 205  
 Lys Asp Glu Tyr Gln Pro Lys Gly Gln Phe Thr Lys Ala Lys Ser Val  
 210 215 220  
 Leu Ala Ile His Asn Ile Ala Phe Gln Gly Arg Met Trp Glu Glu Ala  
 225 230 235 240  
 Phe Lys Asp Thr Lys Leu Pro Gln Ala Ala Phe Asp Lys Leu Ala Phe  
 245 250 255  
 Ser Asp Gly Tyr Ala Lys Val Tyr Thr Glu Ala Thr Pro Met Glu Glu  
 260 265 270  
 Asp Glu Lys Pro Pro Leu Thr Gly Lys Thr Tyr Lys Lys Ile Asn Trp  
 275 280 285  
 Leu Lys Gly Gly Ile Ile Ala Ala Asp Lys Leu Val Thr Val Ser Pro  
 290 295 300  
 Asn Tyr Ala Thr Glu Ile Ala Ala Asp Ala Ala Gly Gly Val Glu Leu  
 305 310 315 320  
 Asp Thr Val Ile Arg Ala Lys Gly Ile Glu Gly Ile Val Asp Gly Met  
 325 330 335  
 Asp Ile Glu Glu Trp Asn Pro Lys Thr Asp Lys Phe Leu Ser Ala Pro  
 340 345 350  
 Tyr Asp Gln Asn Ser Val Tyr Ala Gly Lys Ala Ala Ala Lys Glu Ala  
 355 360 365  
 Leu Gln Ala Glu Leu Gly Leu Pro Val Asp Pro Thr Ala Pro Leu Phe  
 370 375 380  
 Ala Phe Ile Gly Arg Leu Glu Glu Gln Lys Gly Val Asp Ile Ile Leu  
 385 390 395 400  
 Ala Ala Leu Pro Lys Ile Leu Ala Thr Pro Lys Val Gln Ile Ala Ile  
 405 410 415

100 105 110  
 115 120 125  
 130 135 140  
 145 150 155 160  
 165 170 175  
 180 185 190  
 195 200 205  
 210 215 220  
 225 230 235 240  
 245 250 255  
 260 265 270  
 275 280 285  
 290 295 300  
 305 310 315 320  
 325 330 335  
 340 345 350  
 355 360 365  
 370 375 380  
 385 390 395 400  
 405 410 415

Leu Gly Thr Gly Lys Ala Ala Tyr Glu Lys Leu Val Asn Ala Ile Gly  
 420 425 430  
 Thr Lys Tyr Lys Gly Arg Ala Lys Gly Val Val Lys Phe Ser Ala Pro  
 435 440 445  
 Leu Ala His Met Leu Thr Ala Gly Ala Asp Phe Met Leu Val Pro Ser  
 450 455 460  
 Arg Phe Glu Pro Cys Gly Leu Ile Gln Leu His Ala Met His Tyr Gly  
 465 470 475 480  
 Thr Val Pro Val Val Ala Ser Thr Gly Gly Leu Val Asp Thr Val Lys  
 485 490 495  
 Glu Gly Val Thr Gly Phe His Met Gly Ala Leu Asn Pro Asp Lys Leu  
 500 505 510  
 Asp Glu Ala Asp Ala Asp Ala Leu Ala Ala Thr Val Arg Arg Ala Ser  
 515 520 525  
 Glu Val Phe Ala Gly Gly Arg Tyr Pro Glu Met Val Ala Asn Cys Ile  
 530 535 540  
 Ser Gln Asp Leu Ser Trp Ser Lys Pro Ala Gln Lys Trp Glu Gly Leu  
 545 550 555 560  
 Leu Glu Glu Val Val Tyr Gly Lys Gly Gly Val Ala Thr Ala Lys Lys  
 565 570 575  
 Glu Glu Ile Lys Val Pro Val Ala Glu Lys Ile Pro Gly Asp Leu Pro  
 580 585 590  
 Ala Val Ser Tyr Ala Pro Asn Thr Leu Lys Pro Val Ser Ala Ser Val  
 595 600 605  
 Glu Gly Asn Gly Ala Ala Ala Pro Lys Val Gly Thr Thr Ala Pro Ala  
 610 615 620  
 Met Gly Ala Trp Arg Ala Thr Thr Pro Ser Gly Pro Ser Pro Ala Ala  
 625 630 635 640  
 Ala Thr Pro Lys Val Thr Thr Tyr Lys Pro Ala Leu Pro Ala Thr Ala  
 645 650 655  
 Lys Pro Lys Thr Ala Gly Leu Lys Leu Ala Gly Glu Ala Ser Thr Thr  
 660 665 670  
 Ser Thr Ser Glu Asn Gly Ala Ala Ser Asn Gly Asn Gly Asn Gly Ala  
 675 680 685  
 Ser Ala Ser Lys Thr Ser Ala Ala Lys Pro Leu Val Ser Ala Ala Thr  
 690 695 700  
 Arg Lys Ser Ala  
 705

<210> 4  
 <211> 1953  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Description of the artificial sequence: fragment  
 of the complete cDNA coding for the GBSSI of  
 Chlamydomonas reinhardtii and coding for the mature GBSSI  
 protein.

<220>  
 <221> CDS  
 <222> (1)..(1953)

<400> 4  
 gcg ctg gac atc gtg atg gtt gct gct gag gtc gcc cct tgg tcc aag 48  
 Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys  
 1 5 10 15  
 acg ggc ggc ctg ggc gat gtg act ggt ggc ctg cct att gag ctg gtc 96  
 Thr Gly Gly Leu Gly Asp Val Thr Gly Gly Leu Pro Ile Glu Leu Val  
 20 25 30  
 aag cgc ggc cac cgc gtc atg acc att gcc cct cgc tac gac cag tac 144  
 Lys Arg Gly His Arg Val Met Thr Ile Ala Pro Arg Tyr Asp Gln Tyr  
 35 40 45  
 gct gac gcc tgg gac acc tcg gtg gtc gtg gac atc atg ggc gag aag 192  
 Ala Asp Ala Trp Asp Thr Ser Val Val Val Asp Ile Met Gly Glu Lys  
 50 55 60  
 gtc cgc tac ttc cac tcc atc aag aag ggc gtg cac cgc gtg tgg att 240  
 Val Arg Tyr Phe His Ser Ile Lys Lys Gly Val His Arg Val Trp Ile  
 65 70 75 80  
 gac cac ccc tgg ttc ctg gcc aag gtc tgg ggc aag acc ggc tcc aag 288  
 Asp His Pro Trp Phe Leu Ala Lys Val Trp Gly Lys Thr Gly Ser Lys  
 85 90 95  
 ctg tac ggc ccc cgc tcc ggc gct gac tac ctg gac aac cac aag cgc 336  
 Leu Tyr Gly Pro Arg Ser Gly Ala Asp Tyr Leu Asp Asn His Lys Arg  
 100 105 110  
 ttc gcc ctg ttc tgc aag gcc gct att gag gct gcc cgc gtg ctg ccc 384  
 Phe Ala Leu Phe Cys Lys Ala Ala Ile Glu Ala Ala Arg Val Leu Pro  
 115 120 125  
 ttc ggc ccc ggc gag gac tgc gtc ttc gtg gcc aac gac tgg cac tcc 432  
 Phe Gly Pro Gly Glu Asp Cys Val Phe Val Ala Asn Asp Trp His Ser  
 130 135 140  
 gcc ctg gtg ccc gtc ctg ctg aag gac gag tac cag ccc aag ggc cag 480  
 Ala Leu Val Pro Val Leu Leu Lys Asp Glu Tyr Gln Pro Lys Gly Gln  
 145 150 155 160  
 ttc acc aag gcc aag tcg gtg ctg gct atc cac aac atc gcc ttc cag 528  
 Phe Thr Lys Ala Lys Ser Val Leu Ala Ile His Asn Ile Ala Phe Gln  
 165 170 175



ggc cgc atg tgg gag gag gct ttc aag gac acg aag ctg ccc cag gcc	576
Gly Arg Met Trp Glu Glu Ala Phe Lys Asp Thr Lys Leu Pro Gln Ala	
180 185 190	
gcc ttt gac aag ctg gcc ttc tcg gac ggc tat gcc aag gtt tac act	624
Ala Phe Asp Lys Leu Ala Phe Ser Asp Gly Tyr Ala Lys Val Tyr Thr	
195 200 205	
gag gcc acc ccc atg gag gag gac gag aag ccc ccg ctg acg gga aag	672
Glu Ala Thr Pro Met Glu Glu Asp Glu Lys Pro Pro Leu Thr Gly Lys	
210 215 220	
acc tac aag aag atc aac tgg ctg aag ggt ggc att atc gcc gcc gac	720
Thr Tyr Lys Lys Ile Asn Trp Leu Lys Gly Gly Ile Ile Ala Ala Asp	
225 230 235 240	
aag ctg gtg act gtg tcg ccc aac tac gcg acc gag atc gct gcc gat	768
Lys Leu Val Thr Val Ser Pro Asn Tyr Ala Thr Glu Ile Ala Ala Asp	
245 250 255	
gcc gcc ggc ggt gtg gag ctg gac acc gtc atc cgc gcc aag ggc att	816
Ala Ala Gly Gly Val Glu Leu Asp Thr Val Ile Arg Ala Lys Gly Ile	
260 265 270	
gag ggc att gtg aac ggc atg gac att gag gag tgg aac ccc aag acc	864
Glu Gly Ile Val Asn Gly Met Asp Ile Glu Glu Trp Asn Pro Lys Thr	
275 280 285	
gac aag ttc ctg tct gcg ccc tac gac cag aac agc gtc tac gcc ggc	912
Asp Lys Phe Leu Ser Ala Pro Tyr Asp Gln Asn Ser Val Tyr Ala Gly	
290 295 300	
aag gcc gcc gcc aag gag gcc ctg cag gcc gag ctg ggc ctg cct gtg	960
Lys Ala Ala Ala Lys Glu Ala Leu Gln Ala Glu Leu Gly Leu Pro Val	
305 310 315 320	
gac ccc acc gcc ccc ctg ttc gcc ttc atc ggc cgc ctg gag gag cag	1008
Asp Pro Thr Ala Pro Leu Phe Ala Phe Ile Gly Arg Leu Glu Glu Gln	
325 330 335	
aag ggt gtg gac atc atc ctg gcc gcc ctg ccc aag atc ctg gcc acc	1056
Lys Gly Val Asp Ile Ile Leu Ala Ala Leu Pro Lys Ile Leu Ala Thr	
340 345 350	
ccc aag gtg cag atc gcc atc ctg ggt acc ggc aag gcc gcc tac gag	1104
Pro Lys Val Gln Ile Ala Ile Leu Gly Thr Gly Lys Ala Ala Tyr Glu	
355 360 365	
aag ctg gtg aac gcc atc ggc acc aag tac aag ggc cgc gcc aag ggc	1152
Lys Leu Val Asn Ala Ile Gly Thr Lys Tyr Lys Gly Arg Ala Lys Gly	
370 375 380	
gtg gtc aag ttc tcg gcg ccc ctg gcg cac atg ctc acc gcc ggc gcc	1200
Val Val Lys Phe Ser Ala Pro Leu Ala His Met Leu Thr Ala Gly Ala	
385 390 395 400	
gac ttc atg ctg gtg ccc tcg cgc ttc gag ccc tgc ggc ctg atc cag	1248
Asp Phe Met Leu Val Pro Ser Arg Phe Glu Pro Cys Gly Leu Ile Gln	
405 410 415	

ctg	cac	gcc	atg	cac	tac	ggt	acc	gtg	ccc	gtg	gta	gcc	tcc	acc	ggc	1296
Leu	His	Ala	Met	His	Tyr	Gly	Thr	Val	Pro	Val	Val	Ala	Ser	Thr	Gly	
		420						425					430			
ggc	ctg	gtc	gac	acc	gtc	aag	gag	ggc	gtc	acc	ggc	ttc	cac	atg	ggc	1344
Gly	Leu	Val	Asp	Thr	Val	Lys	Glu	Gly	Val	Thr	Gly	Phe	His	Met	Gly	
		435					440					445				
gcc	ctg	aac	ccc	gac	aag	ctg	gac	gag	gct	gac	gcc	gac	gcc	ctg	gcc	1392
Ala	Leu	Asn	Pro	Asp	Lys	Leu	Asp	Glu	Ala	Asp	Ala	Asp	Ala	Leu	Ala	
		450				455					460					
gcc	acc	gtg	cgc	cgt	gcc	agc	gag	gtg	ttt	gcg	ggc	ggc	cgc	tac	ccc	1440
Ala	Thr	Val	Arg	Arg	Ala	Ser	Glu	Val	Phe	Ala	Gly	Gly	Arg	Tyr	Pro	
		465			470				475						480	
gag	atg	gtg	gcc	aac	tgc	atc	agc	cag	gac	ctg	tcc	tgg	tcc	aag	ccc	1488
Glu	Met	Val	Ala	Asn	Cys	Ile	Ser	Gln	Asp	Leu	Ser	Trp	Ser	Lys	Pro	
				485				490						495		
gcc	cag	aag	tgg	gag	ggc	ctg	ctg	gag	gag	gtg	gtg	tac	ggc	aag	ggc	1536
Ala	Gln	Lys	Trp	Glu	Gly	Leu	Leu	Glu	Glu	Val	Val	Tyr	Gly	Lys	Gly	
			500					505					510			
ggc	gtg	gcc	acc	gcc	aag	aag	gag	gag	atc	aag	gtg	ccc	gtt	gcc	gag	1584
Gly	Val	Ala	Thr	Ala	Lys	Lys	Glu	Glu	Ile	Lys	Val	Pro	Val	Ala	Glu	
		515					520					525				
aag	atc	ccc	ggc	gac	ctg	ccc	gcc	gtg	tcc	tac	gcc	ccc	aac	acc	ctg	1632
Lys	Ile	Pro	Gly	Asp	Leu	Pro	Ala	Val	Ser	Tyr	Ala	Pro	Asn	Thr	Leu	
		530				535					540					
aag	ccc	gtg	tcc	gcc	tcc	gtg	gag	ggc	aac	ggc	gcc	gcc	gcg	ccc	aag	1680
Lys	Pro	Val	Ser	Ala	Ser	Val	Glu	Gly	Asn	Gly	Ala	Ala	Ala	Pro	Lys	
		545			550				555						560	
gtc	ggc	acc	acc	gcc	ccc	gcc	atg	ggc	gcg	tgg	cgc	gcg	acc	acc	ccc	1728
Val	Gly	Thr	Thr	Ala	Pro	Ala	Met	Gly	Ala	Trp	Arg	Ala	Thr	Thr	Pro	
				565				570						575		
tcg	ggc	ccc	tcg	ccc	gcc	gcc	gcc	acc	ccc	aag	gtg	acc	acc	tac	aag	1776
Ser	Gly	Pro	Ser	Pro	Ala	Ala	Ala	Thr	Pro	Lys	Val	Thr	Thr	Tyr	Lys	
			580					585						590		
ccc	gcc	ctg	ccc	gcc	acc	gcc	aag	ccc	aag	acc	gct	ggc	ctc	aag	ctg	1824
Pro	Ala	Leu	Pro	Ala	Thr	Ala	Lys	Pro	Lys	Thr	Ala	Gly	Leu	Lys	Leu	
		595					600					605				
gcc	ggt	gag	gcc	tcc	acc	acc	tcg	acc	tcg	gag	aac	ggc	gct	gcc	tcc	1872
Ala	Gly	Glu	Ala	Ser	Thr	Thr	Ser	Thr	Ser	Glu	Asn	Gly	Ala	Ala	Ser	
		610				615					620					
aac	ggc	aac	ggc	aac	ggt	gcc	tcg	gcc	tcc	aag	acc	tcg	gct	gcc	aag	1920
Asn	Gly	Asn	Gly	Asn	Gly	Ala	Ser	Ala	Ser	Lys	Thr	Ser	Ala	Ala	Lys	
		625			630					635					640	
ccc	ctg	gtc	tcc	gcc	gcc	acc	cgc	aag	tcc	gcc						1953
Pro	Leu	Val	Ser	Ala	Ala	Thr	Arg	Lys	Ser	Ala						
				645					650							

<210> 5  
 <211> 651  
 <212> PRT  
 <213> Artificial sequence  
 <223> Description of the artificial sequence: fragment  
 of the complete cDNA coding for the GBSSI  
 of *Chlamydomonas reinhardtii* and coding for the mature  
 GBSSI protein

<400> 5  
 Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys  
 1 5 10 15  
 Thr Gly Gly Leu Gly Asp Val Thr Gly Gly Leu Pro Ile Glu Leu Val  
 20 25 30  
 Lys Arg Gly His Arg Val Met Thr Ile Ala Pro Arg Tyr Asp Gln Tyr  
 35 40 45  
 Ala Asp Ala Trp Asp Thr Ser Val Val Val Asp Ile Met Gly Glu Lys  
 50 55 60  
 Val Arg Tyr Phe His Ser Ile Lys Lys Gly Val His Arg Val Trp Ile  
 65 70 75 80  
 Asp His Pro Trp Phe Leu Ala Lys Val Trp Gly Lys Thr Gly Ser Lys  
 85 90 95  
 Leu Tyr Gly Pro Arg Ser Gly Ala Asp Tyr Leu Asp Asn His Lys Arg  
 100 105 110  
 Phe Ala Leu Phe Cys Lys Ala Ala Ile Glu Ala Ala Arg Val Leu Pro  
 115 120 125  
 Phe Gly Pro Gly Glu Asp Cys Val Phe Val Ala Asn Asp Trp His Ser  
 130 135 140  
 Ala Leu Val Pro Val Leu Leu Lys Asp Glu Tyr Gln Pro Lys Gly Gln  
 145 150 155 160  
 Phe Thr Lys Ala Lys Ser Val Leu Ala Ile His Asn Ile Ala Phe Gln  
 165 170 175  
 Gly Arg Met Trp Glu Glu Ala Phe Lys Asp Thr Lys Leu Pro Gln Ala  
 180 185 190  
 Ala Phe Asp Lys Leu Ala Phe Ser Asp Gly Tyr Ala Lys Val Tyr Thr  
 195 200 205  
 Glu Ala Thr Pro Met Glu Glu Asp Glu Lys Pro Pro Leu Thr Gly Lys  
 210 215 220  
 Thr Tyr Lys Lys Ile Asn Trp Leu Lys Gly Gly Ile Ile Ala Ala Asp  
 225 230 235 240  
 Lys Leu Val Thr Val Ser Pro Asn Tyr Ala Thr Glu Ile Ala Ala Asp  
 245 250 255  
 Ala Ala Gly Gly Val Glu Leu Asp Thr Val Ile Arg Ala Lys Gly Ile  
 260 265 270



<400> 6																
gcg	ctg	gac	atc	gtg	atg	gtt	gct	gct	gag	gtc	gcc	cct	tgg	tcc	aag	48
Ala	Leu	Asp	Ile	Val	Met	Val	Ala	Ala	Glu	Val	Ala	Pro	Trp	Ser	Lys	
1				5					10					15		
acg	ggc	ggc	ctg	ggc	gat	gtg	act	ggt	ggc	ctg	cct	att	gag	ctg	gtc	96
Thr	Gly	Gly	Leu	Gly	Asp	Val	Thr	Gly	Gly	Leu	Pro	Ile	Glu	Leu	Val	
			20					25					30			
aag	cgc	ggc	cac	cgc	gtc	atg	acc	att	gcc	cct	cgc	tac	gac	cag	tac	144
Lys	Arg	Gly	His	Arg	Val	Met	Thr	Ile	Ala	Pro	Arg	Tyr	Asp	Gln	Tyr	
		35					40					45				
gct	gac	gcc	tgg	gac	acc	tcg	gtg	gtc	gtg	gac	atc	atg	ggc	gag	aag	192
Ala	Asp	Ala	Trp	Asp	Thr	Ser	Val	Val	Val	Asp	Ile	Met	Gly	Glu	Lys	
	50					55					60					
gtc	cgc	tac	ttc	cac	tcc	atc	aag	aag	ggc	gtg	cac	cgc	gtg	tgg	att	240
Val	Arg	Tyr	Phe	His	Ser	Ile	Lys	Lys	Gly	Val	His	Arg	Val	Trp	Ile	
65					70				75					80		
gac	cac	ccc	tgg	ttc	ctg	gcc	aag	gtc	tgg	ggc	aag	acc	ggc	tcc	aag	288
Asp	His	Pro	Trp	Phe	Leu	Ala	Lys	Val	Trp	Gly	Lys	Thr	Gly	Ser	Lys	
				85				90						95		
ctg	tac	ggc	ccc	cgc	tcc	ggc	gct	gac	tac	ctg	gac	aac	cac	aag	cgc	336
Leu	Tyr	Gly	Pro	Arg	Ser	Gly	Ala	Asp	Tyr	Leu	Asp	Asn	His	Lys	Arg	
			100					105					110			
ttc	gcc	ctg	ttc	tgc	aag	gcc	gct	att	gag	gct	gcc	cgc	gtg	ctg	ccc	384
Phe	Ala	Leu	Phe	Cys	Lys	Ala	Ala	Ile	Glu	Ala	Ala	Arg	Val	Leu	Pro	

115	120	125	
ttc ggc ccc ggc gag gac tgc gtc ttc gtg gcc aac gac tgg cac tcc Phe Gly Pro Gly Glu Asp Cys Val Phe Val Ala Asn Asp Trp His Ser 130 135 140			432
gcc ctg gtg ccc gtc ctg ctg aag gac gag tac cag ccc aag ggc cag Ala Leu Val Pro Val Leu Lys Asp Glu Tyr Gln Pro Lys Gly Gln 145 150 155 160			480
ttc acc aag gcc aag tcg gtg ctg gct atc cac aac atc gcc ttc cag Phe Thr Lys Ala Lys Ser Val Leu Ala Ile His Asn Ile Ala Phe Gln 165 170 175			528
ggc cgc atg tgg gag gag gct ttc aag gac acg aag ctg ccc cag gcc Gly Arg Met Trp Glu Glu Ala Phe Lys Asp Thr Lys Leu Pro Gln Ala 180 185 190			576
gcc ttt gac aag ctg gcc ttc tcg gac ggc tat gcc aag gtt tac act Ala Phe Asp Lys Leu Ala Phe Ser Asp Gly Tyr Ala Lys Val Tyr Thr 195 200 205			624
gag gcc acc ccc atg gag gag gac gag aag ccc ccg ctg acg gga aag Glu Ala Thr Pro Met Glu Glu Asp Glu Lys Pro Pro Leu Thr Gly Lys 210 215 220			672
acc tac aag aag atc aac tgg ctg aag ggt ggc att atc gcc gcc gac Thr Tyr Lys Lys Ile Asn Trp Leu Lys Gly Gly Ile Ile Ala Ala Asp 225 230 235 240			720
aag ctg gtg act gtg tcg ccc aac tac gcg acc gag atc gct gcc gat Lys Leu Val Thr Val Ser Pro Asn Tyr Ala Thr Glu Ile Ala Ala Asp 245 250 255			768
gcc gcc ggc ggt gtg gag ctg gac acc gtc atc cgc gcc aag ggc att Ala Ala Gly Gly Val Glu Leu Asp Thr Val Ile Arg Ala Lys Gly Ile 260 265 270			816
gag ggc att gtg aac ggc atg gac att gag gag tgg aac ccc aag acc Glu Gly Ile Val Asn Gly Met Asp Ile Glu Glu Trp Asn Pro Lys Thr 275 280 285			864
gac aag ttc ctg tct gcg ccc tac gac cag aac agc gtc tac gcc ggc Asp Lys Phe Leu Ser Ala Pro Tyr Asp Gln Asn Ser Val Tyr Ala Gly 290 295 300			912
aag gcc gcc gcc aag gag gcc ctg cag gcc gag ctg ggc ctg cct gtg Lys Ala Ala Ala Lys Glu Ala Leu Gln Ala Glu Leu Gly Leu Pro Val 305 310 315 320			960
gac ccc acc gcc ccc ctg ttc gcc ttc atc ggc cgc ctg gag gag cag Asp Pro Thr Ala Pro Leu Phe Ala Phe Ile Gly Arg Leu Glu Glu Gln 325 330 335			1008
aag ggt gtg gac atc atc ctg gcc gcc ctg ccc aag atc ctg gcc acc Lys Gly Val Asp Ile Ile Leu Ala Ala Leu Pro Lys Ile Leu Ala Thr 340 345 350			1056
ccc aag gtg cag atc gcc atc ctg ggt acc ggc aag gcc gcc tac gag Pro Lys Val Gln Ile Ala Ile Leu Gly Thr Gly Lys Ala Ala Tyr Glu 355 360 365			1104

aag ctg gtg aac gcc atc ggc acc aag tac aag ggc cgc gcc aag ggc 1152  
 Lys Leu Val Asn Ala Ile Gly Thr Lys Tyr Lys Gly Arg Ala Lys Gly  
 370 375 380

gtg gtc aag ttc tcg gcg ccc ctg gcg cac atg ctc acc gcc ggc gcc 1200  
 Val Val Lys Phe Ser Ala Pro Leu Ala His Met Leu Thr Ala Gly Ala  
 385 390 395 400

gac ttc atg ctg gtg ccc tcg cgc ttc gag ccc tgc ggc ctg atc cag 1248  
 Asp Phe Met Leu Val Pro Ser Arg Phe Glu Pro Cys Gly Leu Ile Gln  
 405 410 415

ctg cac gcc atg cac tac ggt acc gtg ccc gtg gta gcc tcc acc ggc 1296  
 Leu His Ala Met His Tyr Gly Thr Val Pro Val Val Ala Ser Thr Gly  
 420 425 430

ggc ctg gtc gac acc gtc 1314  
 Gly Leu Val Asp Thr Val  
 435

<210> 7

<211> 438

<212> PRT

<213> Artificial sequence

<223> Description of the artificial sequence: fragment  
 of the complete cDNA coding for the GBSSI  
 of Chlamydomonas reinhardtii

<400> 7

Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys  
 1 5 10 15

Thr Gly Gly Leu Gly Asp Val Thr Gly Gly Leu Pro Ile Glu Leu Val  
 20 25 30

Lys Arg Gly His Arg Val Met Thr Ile Ala Pro Arg Tyr Asp Gln Tyr  
 35 40 45

Ala Asp Ala Trp Asp Thr Ser Val Val Val Asp Ile Met Gly Glu Lys  
 50 55 60

Val Arg Tyr Phe His Ser Ile Lys Lys Gly Val His Arg Val Trp Ile  
 65 70 75 80

Asp His Pro Trp Phe Leu Ala Lys Val Trp Gly Lys Thr Gly Ser Lys  
 85 90 95

Leu Tyr Gly Pro Arg Ser Gly Ala Asp Tyr Leu Asp Asn His Lys Arg  
 100 105 110

Phe Ala Leu Phe Cys Lys Ala Ala Ile Glu Ala Ala Arg Val Leu Pro  
 115 120 125

Phe Gly Pro Gly Glu Asp Cys Val Phe Val Ala Asn Asp Trp His Ser  
 130 135 140

Ala Leu Val Pro Val Leu Leu Lys Asp Glu Tyr Gln Pro Lys Gly Gln  
 145 150 155 160

Phe Thr Lys Ala Lys Ser Val Leu Ala Ile His Asn Ile Ala Phe Gln  
 165 170 175  
 Gly Arg Met Trp Glu Glu Ala Phe Lys Asp Thr Lys Leu Pro Gln Ala  
 180 185 190  
 Ala Phe Asp Lys Leu Ala Phe Ser Asp Gly Tyr Ala Lys Val Tyr Thr  
 195 200 205  
 Glu Ala Thr Pro Met Glu Glu Asp Glu Lys Pro Pro Leu Thr Gly Lys  
 210 215 220  
 Thr Tyr Lys Lys Ile Asn Trp Leu Lys Gly Gly Ile Ile Ala Ala Asp  
 225 230 235 240  
 Lys Leu Val Thr Val Ser Pro Asn Tyr Ala Thr Glu Ile Ala Ala Asp  
 245 250 255  
 Ala Ala Gly Gly Val Glu Leu Asp Thr Val Ile Arg Ala Lys Gly Ile  
 260 265 270  
 Glu Gly Ile Val Asn Gly Met Asp Ile Glu Glu Trp Asn Pro Lys Thr  
 275 280 285  
 Asp Lys Phe Leu Ser Ala Pro Tyr Asp Gln Asn Ser Val Tyr Ala Gly  
 290 295 300  
 Lys Ala Ala Ala Lys Glu Ala Leu Gln Ala Glu Leu Gly Leu Pro Val  
 305 310 315 320  
 Asp Pro Thr Ala Pro Leu Phe Ala Phe Ile Gly Arg Leu Glu Glu Gln  
 325 330 335  
 Lys Gly Val Asp Ile Ile Leu Ala Ala Leu Pro Lys Ile Leu Ala Thr  
 340 345 350  
 Pro Lys Val Gln Ile Ala Ile Leu Gly Thr Gly Lys Ala Ala Tyr Glu  
 355 360 365  
 Lys Leu Val Asn Ala Ile Gly Thr Lys Tyr Lys Gly Arg Ala Lys Gly  
 370 375 380  
 Val Val Lys Phe Ser Ala Pro Leu Ala His Met Leu Thr Ala Gly Ala  
 385 390 395 400  
 Asp Phe Met Leu Val Pro Ser Arg Phe Glu Pro Cys Gly Leu Ile Gln  
 405 410 415  
 Leu His Ala Met His Tyr Gly Thr Val Pro Val Val Ala Ser Thr Gly  
 420 425 430  
 Gly Leu Val Asp Thr Val  
 435

<210> 8  
 <211> 1593  
 <212> DNA



$\langle 220 \rangle$  $\langle 220 \rangle$ 

<222> (1) .. (1593)

gcg ctg gac atc gtg atg gtt gct gct gag gtc gcc cct tgg tcc aag 48  
Ala Leu Asp Ile Val Met Val Ala Ala Glu Val Ala Pro Trp Ser Lys  
1 : 5 10 15

acg ggc ggc ctg ggc gat gtg act ggt ggc ctg cct att gag ctg gtc 96  
Thr Gly Gly Leu Gly Asp Val Thr Gly Gly Leu Pro Ile Glu Leu Val  
20 25 30

aag cgc ggc cac cgc gtc atg acc att gcc cct cgc tac gac cag tac 144  
Lys Arg Gly His Arg Val Met Thr Ile Ala Pro Arg Tyr Asp Gln Tyr  
35 40 45

gct gac gcc tgg gac acc tcg gtg gtc gtg gac atc atg ggc gag aag 192  
Ala Asp Ala Trp Asp Thr Ser Val Val Val Asp Ile Met Gly Glu Lys  
50 55 60

gtc cgc tac ttc cac tcc atc aag aag ggc gtg cac cgc gtg tgg att 240  
Val Arg Tyr Phe His Ser Ile Lys Lys Gly Val His Arg Val Trp Ile  
65 70 75 80

gac cac ccc tgg ttc ctg gcc aag gtc tgg ggc aag acc ggc tcc aag 288  
Asp His Pro Trp Phe Leu Ala Lys Val Trp Gly Lys Thr Gly Ser Lys  
85 90 95

ctg tac ggc ccc cgc tcc ggc gct gac tac ctg gac aac cac aag cgc 336  
Leu Tyr Gly Pro Arg Ser Gly Ala Asp Tyr Leu Asp Asn His Lys Arg  
100 105 110

ttc	gcc	ctg	ttc	tgc	aag	gcc	gct	att	gag	gct	gcc	cgc	gtg	ctg	ccc	384
Phe	Ala	Leu	Phe	Cys	Lys	Ala	Ala	Ile	Glu	Ala	Ala	Arg	Val	Leu	Pro	
		115					120					125				

ttc	ggc	ccc	ggc	gag	gac	tgc	gtc	ttc	gtg	gcc	aac	gac	tgg	cac	tcc	432
Phe	Gly	Pro	Gly	Glu	Asp	Cys	Val	Phe	Val	Ala	Asn	Asp	Trp	His	Ser	
	130					135					140					

gcc ctg gtg ccc gtc ctg ctg aag gac gag tac cag ccc aag ggc cag 480  
Ala Leu Val Pro Val Leu Leu Lys Asp Glu Tyr Gln Pro Lys Gly Gln  
145 150 155 160

ttc acc aag gcc aag tcg gtg ctg gct atc cac aac atc gcc ttc cag    528  
Phe Thr Lys Ala Lys Ser Val Leu Ala Ile His Asn Ile Ala Phe Gln  
                165                    170                    175

ggc cgc atg tgg gag gag gct ttc aag gac acg aag ctg ccc cag gcc 576  
Gly Arg Met Trp Glu Glu Ala Phe Lys Asp Thr Lys Leu Pro Gln Ala  
180 185 190

gcc ttt gac aaq ctg gcc ttc tcg gac ggc tat gcc aaq gtt tac act 624

Ala	Phe	Asp 195	Lys	Leu	Ala	Phe	Ser 200	Asp	Gly	Tyr	Ala	Lys 205	Val	Tyr	Thr	
gag Glu	gcc Ala	acc Thr	ccc Pro	atg Met	gag Glu	gag Glu	gac Asp	gag Glu	aag Lys	ccc Pro	ccg Pro	ctg Leu	acg Thr	gga Gly	aag Lys	672
210						215					220					
acc Thr	tac Tyr	aag Lys	aag Lys	atc Ile	aac Asn	tgg Trp	ctg Leu	aag Lys	ggg Gly	ggc Gly	att Ile	atc Ile	gcc Ala	gcc Ala	gac Asp	720
225					230					235					240	
aag Lys	ctg Leu	gtg Val	act Thr	gtg Val	tcg Ser	ccc Pro	aac Asn	tac Tyr	gcg Ala	acc Thr	gag Glu	atc Ile	gct Ala	gcc Ala	gat Asp	768
245									250					255		
gcc Ala	gcc Ala	ggc Gly	ggg Gly	gtg Val	gag Glu	ctg Leu	gac Asp	acc Thr	gtc Val	atc Ile	cgc Arg	gcc Ala	aag Lys	ggc Gly	att Ile	816
			260					265					270			
gag Glu	ggc Gly	att Ile	gtg Val	aac Asn	ggc Gly	atg Met	gac Asp	att Ile	gag Glu	gag Glu	tgg Trp	aac Asn	ccc Pro	aag Lys	acc Thr	864
	275						280					285				
gac Asp	aag Lys	ttc Phe	ctg Leu	tct Ser	gcg Ala	ccc Pro	tac Tyr	gac Asp	cag Gln	aac Asn	agc Ser	gtc Val	tac Tyr	gcc Ala	ggc Gly	912
	290					295					300					
aag Lys	gcc Ala	gcc Ala	gcc Ala	aag Lys	gag Glu	gcc Ala	ctg Leu	cag Gln	gcc Ala	gag Glu	ctg Leu	ggc Gly	ctg Leu	cct Pro	gtg Val	960
305					310					315					320	
gac Asp	ccc Pro	acc Thr	gcc Ala	ccc Pro	ctg Leu	ttc Phe	gcc Ala	ttc Phe	atc Ile	ggc Gly	cgc Arg	ctg Leu	gag Glu	gag Glu	cag Gln	1008
				325					330					335		
aag Lys	ggg Gly	gtg Val	gac Asp	atc Ile	atc Ile	ctg Leu	gcc Ala	gcc Ala	ctg Leu	ccc Pro	aag Lys	atc Ile	ctg Leu	gcc Ala	acc Thr	1056
			340					345					350			
ccc Pro	aag Lys	gtg Val	cag Gln	atc Ile	gcc Ala	atc Ile	ctg Leu	ggg Gly	acc Thr	ggc Gly	aag Lys	gcc Ala	gcc Ala	tac Tyr	gag Glu	1104
		355					360					365				
aag Lys	ctg Leu	gtg Val	aac Asn	gcc Ala	atc Ile	ggc Gly	acc Thr	aag Lys	tac Tyr	aag Lys	ggc Gly	cgc Arg	gcc Ala	aag Lys	ggc Gly	1152
	370					375					380					
gtg Val	gtc Val	aag Lys	ttc Phe	tcg Ser	gcg Ala	ccc Pro	ctg Leu	gcg Ala	cac His	atg Met	ctc Leu	acc Thr	gcc Ala	ggc Gly	gcc Ala	1200
385					390					395					400	
gac Asp	ttc Phe	atg Met	ctg Leu	gtg Val	ccc Pro	tcg Ser	cgc Arg	ttc Phe	gag Glu	ccc Pro	tcg Cys	ggc Gly	ctg Leu	atc Ile	cag Gln	1248
				405					410					415		
ctg Leu	cac His	gcc Ala	atg Met	cac His	tac Tyr	ggg Gly	acc Thr	gtg Val	ccc Pro	gtg Val	gta Val	gcc Ala	tcc Ser	acc Thr	ggc Gly	1296
			420					425					430			
ggc Gly	ctg Leu	gtc Val	gac Asp	acc Thr	gtc Val	aag Lys	gag Glu	ggc Gly	gtc Val	acc Thr	ggc Gly	ttc Phe	cac His	atg Met	ggc Gly	1344



130	135	140
Ala Leu Val Pro Val	Leu Leu Lys Asp Glu Tyr	Gln Pro Lys Gly Gln
145	150	155 160
Phe Thr Lys Ala Lys	Ser Val Leu Ala Ile His	Asn Ile Ala Phe Gln
	165	170 175
Gly Arg Met Trp Glu	Glu Ala Phe Lys Asp Thr	Lys Leu Pro Gln Ala
	180	185 190
Ala Phe Asp Lys Leu	Ala Phe Ser Asp Gly Tyr	Ala Lys Val Tyr Thr
	195	200 205
Glu Ala Thr Pro Met	Glu Glu Asp Glu Lys Pro	Pro Leu Thr Gly Lys
	210	215 220
Thr Tyr Lys Lys Ile	Asn Trp Leu Lys Gly Gly	Ile Ile Ala Ala Asp
	225	230 235 240
Lys Leu Val Thr Val	Ser Pro Asn Tyr Ala Thr	Glu Ile Ala Ala Asp
	245	250 255
Ala Ala Gly Gly Val	Glu Leu Asp Thr Val Ile	Arg Ala Lys Gly Ile
	260	265 270
Glu Gly Ile Val Asn	Gly Met Asp Ile Glu Glu	Trp Asn Pro Lys Thr
	275	280 285
Asp Lys Phe Leu Ser	Ala Pro Tyr Asp Gln Asn	Ser Val Tyr Ala Gly
	290	295 300
Lys Ala Ala Ala Lys	Glu Ala Leu Gln Ala Glu	Leu Gly Leu Pro Val
	305	310 315 320
Asp Pro Thr Ala Pro	Leu Phe Ala Phe Ile Gly	Arg Leu Glu Glu Gln
	325	330 335
Lys Gly Val Asp Ile	Ile Leu Ala Ala Leu Pro	Lys Ile Leu Ala Thr
	340	345 350
Pro Lys Val Gln Ile	Ala Ile Leu Gly Thr Gly	Lys Ala Ala Tyr Glu
	355	360 365
Lys Leu Val Asn Ala	Ile Gly Thr Lys Tyr Lys	Gly Arg Ala Lys Gly
	370	375 380
Val Val Lys Phe Ser	Ala Pro Leu Ala His Met	Leu Thr Ala Gly Ala
	385	390 395 400
Asp Phe Met Leu Val	Pro Ser Arg Phe Glu Pro	Cys Gly Leu Ile Gln
	405	410 415
Leu His Ala Met His	Tyr Gly Thr Val Pro Val	Val Ala Ser Thr Gly
	420	425 430
Gly Leu Val Asp Thr	Val Lys Glu Gly Val Thr	Gly Phe His Met Gly
	435	440 445
Ala Leu Asn Pro Asp	Lys Leu Asp Glu Ala Asp	Ala Asp Ala Leu Ala
	450	455 460

130 135 140

Lys Ile Pro  
530

[illegible]